

The Construction and Application Exploration of Medical Full-text Knowledge Extraction System based on Deep Learning

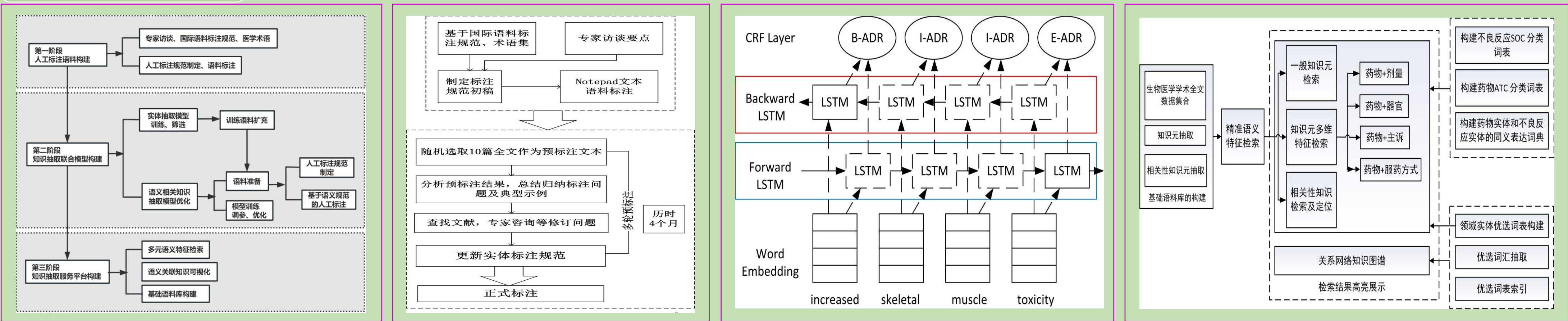
Yuan,xiaoyuan Daijuan yuanxy@nju.edu.cn
Jiangsu, Nanjing, Nanjing University Library



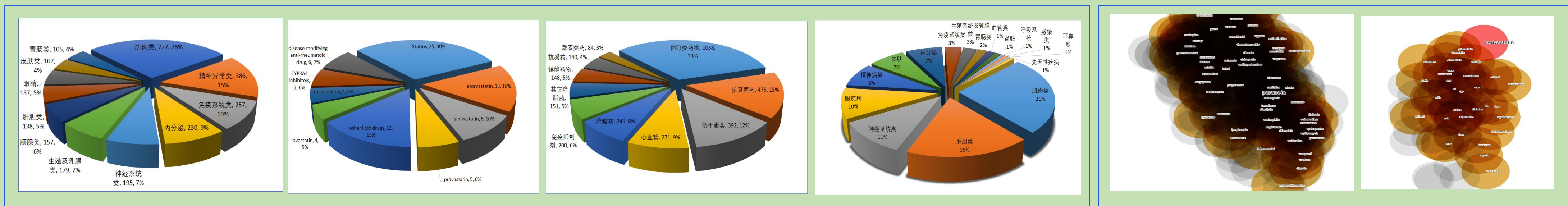
Background

In recent years, the global volume of medical literature has experienced explosive growth. How to efficiently extract precise knowledge from massive medical literature and build intelligent knowledge service systems has remained a core challenge in the library field. This study aims to combine deep learning technologies to design efficient medical Named Entity Recognition (NER) and relationship extraction service frameworks addressing the fine-grained knowledge needs of medical researchers, while validating their effectiveness in practical application scenarios.

Work Flow



Results and Discussion



Conclusion

- The experiment constructs a manually annotated corpus of statin-ADR relationships, followed by model training, optimization, and data augmentation.
- The Bi-LSTM-CRF and Bi-LSTM architectures demonstrate superior performance: drug entity detection achieves F1=0.9, ADR detection F1=0.82, and Bi-LSTM outperforms attention-based At-Bi-LSTM and traditional SVM methods in drug-ADR relation extraction (F1=0.8).
- Leveraging 1,602 annotated full-text articles, we developed a domain knowledge extraction platform (<http://lk.fmcib.com.cn:19081/nlp/>) featuring intelligent reading paradigms for medical researchers.

Application <http://lk.fmcib.com.cn:19081/nlp/>



Reference

Hang Wang, et al. Causal-relationship representation enhanced joint extraction model for elements and relationships[J]. Neurocomputing, 2025,613(1).
Xiaocheng Luo, et al. A bi-consolidating model for joint relational triple extraction[J], Neurocomputing, 2025,614(21):128768.
Jiandong Liu, A cascaded retrieval-while-reasoning multi-document comprehension framework with incremental attention for medical question answering[J], Expert Systems with Applications, 2025,265(15):125701.
Jinbin Zhang, More intelligent knowledge graph: A large language model-driven method for knowledge presentation in geospatial digital twins[J], International Journal of Applied Earth Observation and Geoinformation, 2025, 139:104527.